Artificial intelligence in healthcare: Proposal for a new medico-legal methodology in medical liability

Legal Medicine 80 (2026) 102764

Contents lists available at ScienceDirect

Legal Medicine

journal homepage: www.elsevier.com/locate/legalmed

Artificial intelligence in healthcare: Proposal for a new medico-legal methodology in medical liability

Rossana Cecchi [a], Francesco Calabrò [b], Jessika Camatti [b,*], Anna Laura Santunione [c], Michela Sperti [d], Eric Adriano Zizzi [d], Marco Agostino Deriu [d]

[a] University of Modena and Reggio Emilia, Italy
[b] University of Parma, Italy
[c] University of Modena and Reggio Emilia, Modena, Italy
[d] Polito (BIO)Med Lab Torino, Italy

ARTICLE INFO

Keywords:
Artificial Intelligence
AI
Legal Medicine
Medical Liability
Malpractice

ABSTRACT

The rapid integration of Artificial Intelligence (AI) into healthcare promises significant benefits but also raises unprecedented ethical, clinical, and legal challenges. Current medico-legal frameworks, primarily designed for human decision-making, are often inadequate to address liability issues arising from algorithmic errors or opaque "black box" models. This paper introduces a novel medico-legal methodology that combines proactive and reactive approaches to risk assessment, originally developed within European forensic medicine, and adapts it to the context of AI in healthcare. By systematically analyzing data collection, dataset validation, error identification, and causal reconstruction, the proposed framework provides a structured path for evaluating medical liability when AI systems are involved. This dual approach not only supports clinicians, developers, and policymakers in preventing harm, but also establishes a robust forensic tool for liability assessment. The methodology offers a step toward internationally applicable standards for addressing the medico-legal implications of AI in medicine.

1. Introduction

Over the past decades, the adoption of innovative medical technologies has often followed a recurring pattern: initial enthusiasm, subsequent disappointment, gradual consolidation, and eventual routine use—a process widely described as the "hype cycle" [1]. Artificial Intelligence (AI) in healthcare appears to be following the same trajectory. Current applications generate remarkable expectations, but also expose critical challenges, particularly in terms of safety, transparency, and liability [2–5].

Unlike traditional medical tools, AI systems are characterized by rapid algorithmic evolution, potential biases in training data, and the opacity of complex "black box" models. These features amplify uncertainty and make it difficult to apply existing medico-legal standards, which were developed to evaluate human decision-making. Predicting long-term consequences of AI deployment, both positive and negative, is an additional challenge that existing regulatory frameworks only partially address [6,7].

In this context, there is a pressing need for methodologies that can guide clinicians, developers, and policymakers in balancing technological innovation with patient safety and accountability [8,9]. Forensic medicine offers a valuable perspective, as it has long combined two complementary approaches: (a) a priori risk assessment aimed at prevention, and (b) a posteriori causal analysis aimed at reconstructing liability after adverse events [6].

Building on the European Council of Legal Medicine (ECLM) Guidelines [10], which represent a widely adopted reference for the evaluation of medical liability, we propose an adaptation of this dual methodology to AI in healthcare. After its initial presentation to the Italian medico-legal community in 2024, we now intend to introduce this adapted methodology to the international scientific community, with the aim of fostering a broader debate on the medico-legal implications of AI in healthcare [11]. Our approach integrates proactive and reactive strategies to systematically assess risks, identify errors, and reconstruct causal chains when AI systems are implicated in patient harm. By applying medico-legal reasoning to algorithmic decision-

* Corresponding author.
E-mail addresses: rossana.cecchi@unimore.it (R. Cecchi), francesco.calabro@unipr.it (F. Calabrò), jessika.camatti@unipr.it (J. Camatti), annalaura.santunione@unimore.it (A.L. Santunione), michela.sperti@polito.it (M. Sperti), eric.zizzi@polito.it (E.A. Zizzi), marco.deriu@polito.it (M.A. Deriu).

https://doi.org/10.1016/j.legalmed.2025.102764
Received 20 October 2025; Received in revised form 21 November 2025; Accepted 9 December 2025
Available online 11 December 2025

making, we aim to contribute to the development of robust, internationally applicable standards for the safe and accountable use of AI in medicine.

2. Materials and methods

This study explores how a medico-legal methodology, originally developed for assessing professional liability in clinical practice, can be adapted to address liability issues arising from the use of AI in healthcare. The proposed methodology derives from the work of an interdisciplinary task force specifically constituted for this study, composed of forensic medicine specialists and engineers with expertise in AI systems. This collaborative structure was established to ensure methodological rigor and technical accuracy in adapting medico-legal tools to the evaluation of AI-related liability.

The reference framework is the 13-step methodology endorsed by the European Council of Legal Medicine (ECLM) [10], which systematically integrates both proactive (preventive) and reactive (retrospective) approaches. To adapt this methodology to AI, we analyzed each of the original 13 steps and identified their correspondence within the lifecycle of AI systems. Particular emphasis was placed on: proactive steps: evaluation of data availability and dataset validation before model development, aimed at ensuring scientific reliability and minimizing bias; transitional steps: identification of incorrect outputs and assessment of the model's explainability; reactive steps: verification of inputs, reassessment of model performance, classification of errors, causal analysis, and damage estimation.

The adaptation process was conducted with the goal of maintaining methodological rigor while addressing the specific challenges of AI technologies, such as data bias, algorithmic opacity, and evolving model performance over time.

The modified framework was then compared step-by-step with the original ECLM methodology, highlighting points of overlap, divergence, and the need for additional expertise (e.g., data scientists, bioengineers) to support the medico-legal evaluation of AI-related adverse events.

Operationally, the adapted 13-step structure is currently implemented as a set of prompts and structured questions that can be used with existing generative AI models (e.g., large language models) to guide the systematic collection and analysis of information in AI-related medico-legal cases. At the same time, these prompts are conceived as a preparatory layer for the future development of dedicated software tools or decision-support interfaces, in which the same logic could be embedded in a more formalised workflow tailored to different healthcare and medico-legal settings.

3. Results

The adaptation of the ECLM methodology to AI in healthcare produced a structured framework that integrates proactive, transitional, and reactive phases of analysis. While the original 13-step methodology was designed to assess liability in clinical practice, its adaptation allows for systematic evaluation of algorithmic systems.

Proactive steps (1–2): These focus on the preliminary phases of AI development, particularly the comparative assessment of available data and the validation of datasets against scientific evidence. This stage is critical for minimizing bias, ensuring dataset representativeness, and verifying the model's initial reliability.

Transitional steps (3–5): These steps address the identification of erroneous outputs and the determination of the correct or expected outputs. A central element introduced in the adaptation is the evaluation of whether the AI system is "explainable" (XAI) or a "black box," as this determines the feasibility of reconstructing the decision-making process.

Reactive steps (6–13): Once an error has been confirmed, the methodology shifts to retrospective analysis. This includes verification of inputs, reassessment of model performance, error classification, causal analysis, and final damage estimation. Importantly, step 13 has been expanded beyond traditional damage assessment to include recommendations for model improvement, integrating risk management principles into the medico-legal evaluation.

Table 1 summarizes the comparison between the original ECLM steps and their adaptation to AI-related contexts. This comparative framework highlights the need to integrate technical expertise (e.g., data scientists, engineers) with medico-legal reasoning to ensure comprehensive evaluation. It also demonstrates how medico-legal methodologies can support both prevention of errors during AI development and liability assessment after adverse events. Further details are reported in Table 2.

4. Discussion

The rapid deployment of AI in healthcare has triggered a global debate on safety, accountability, and the limits of existing medico-legal frameworks. While regulatory initiatives such as the European Union's AI Act [12], the U.S. Food and Drug Administration's guidance for Software as a Medical Device [13], and the World Health Organization's guidelines on ethics and governance of AI for health [14] provide important governance structures, they remain largely normative. What is often missing is a practical medico-legal methodology capable of guiding case-specific evaluations when adverse events occur. Our work addresses this gap by adapting a consolidated forensic tool to the specific challenges posed by AI systems.

A central strength of the proposed framework is its dual structure. By combining proactive risk assessment with reactive causal reconstruction, it mirrors the iterative nature of machine learning itself. In the proactive phase, particular emphasis is placed on dataset integrity, representativeness, and bias minimization, which are crucial determinants of model reliability [15,16]. The reactive phase then allows systematic evaluation of erroneous outputs, reconstruction of decision-making pathways, and quantification of damage. Importantly, the methodology expands beyond classical medico-legal assessment by feeding conclusions back into model improvement, thus linking liability analysis with risk management and patient safety [17].

Another key contribution lies in the explicit integration of explainability into medico-legal reasoning. The ability—or inability—to trace how an algorithm produced a given output has profound consequences for liability attribution. For explainable AI (XAI) models, causal pathways can be reconstructed with relative clarity, whereas black-box systems pose substantial evidentiary challenges [18,19]. This distinction underscores the need for interdisciplinary collaboration, where medico-legal experts work alongside clinicians, data scientists, and engineers to ensure both technical validity and forensic robustness.

The present proposal builds upon previous experiences where structured medico-legal methodologies have been applied to complex contexts. For example, the ECLM on-site inspection form has recently demonstrated its value in ensuring consistency and reproducibility across heterogeneous case series [20–22]. This precedent highlights how a standardized and transparent approach can strengthen medico-legal evaluation in practice, supporting the idea that similar principles may be effectively extended to the assessment of AI-related adverse events.

To illustrate how the methodology can be applied in practice, consider the case of an AI-based triage system used in an emergency department, which classifies a patient with chest pain as "low priority" leading to a delayed medical assessment and subsequent myocardial infarction. In the proactive phase, the adapted framework would require verification of the training data and validation procedures used to develop the triage algorithm, including checks on the representation of patients with atypical symptoms (e.g., women or older adults). In the transitional phase, the erroneous output (low-priority label) would be identified and compared with the expected standard of care (urgent cardiological evaluation). In the reactive phase, the analysis would then focus on input integrity (e.g., whether vital signs and risk factors were correctly entered), model performance at the time of the event, error classification (systemic bias versus isolated malfunction), and causal

Table 1
Comparison between the original European Council of Legal Medicine (ECLM) Guidelines and their adaptation to Artificial Intelligence in healthcare.

Original Guidelines	Guidelines applied to AI
STEP 1 — Comparative evaluation of data The medico-legal expert gathers together all the data from the various ascertainment phases, conducts an initial synthesis according to conceptual area and reaches a comparative final evaluation.	**STEP 1 – Comparative assessment of available AI data** Evaluation of the data available to the IA for dataset construction. The construction of a good dataset implies an integrated and multidisciplinary approach consisting of: – *National and supranational laws* – *Domain experts* – *Data collector* Preliminary phase with respect to the development of the model.
STEP 2 — Identification of pathological features STEP 1 is followed by identification of *pathological features*, subdivided into *initial*, *intermediate* and *final* clinical pictures resulting in restoration to health, death, chronic pathological state or permanent injury. In this reconstruction, the physiopathological pathways revealing the chain of events must be identified and clearly described.	**STEP 2 — Validation of the dataset based on scientific evidence** When step 1 is satisfied, we move on to step 2, which consists of validating the dataset based on *scientific references* (guidelines, best practices, scientific reference literature). These are used to assess the scientific assumptions on which the model is based, such as the statistically significant size of the dataset, the minimization of bias and the performance of the model in terms of specificity and sensitivity. The first two steps are preliminary and independent of the occurrence of an error; subsequent steps are only considered if an erroneous output occurs.
STEP 3 — Damage identification This covers possible damage or incapacity, either temporary or permanent (i.e. death, chronic evolutive disease, sequelae),	**STEP 3 — Identification of incorrect output** Phase 3 considers an AI result that turns out to be erroneous. The aim is to identify it and clarify its nature (e.g. false positive, false negative, error in therapy, incorrect prognosis, error in follow-up, etc.).
STEP 4 — Reconstruction of physiopathological pathways and ideal medical conduct Identified *pathological features* are examined by analyzing *scientific sources*, such as *guidelines* (national and international), *consensus documents* (national and international), *operational procedures* (local, national and international), *evidence-based publications* (Cochrane reviews, *meta*-analyses, etc.) and other *literature* data, composed of treatises and articles published in peer-reviewed journals (PubMed-Medline, Embase, Scopus, Ovid, ISI Web of Science, etc.), preferably with *impact factor*. These scientific sources of non-equivalent importance must also be graduated according to the *source hierarchy* shown below. This examination aims at: (a) identifying and reconstructing the physiopathological course composing the actual chain of events which took place, i.e. linking the initial pathological features with the intermediate and final ones; (b) reconstructing the ideal conduct which a physician should have followed during diagnosis, prognosis and treatment. – Guidelines – Consensus documents – Operational procedures – Evidence-based publications – National literature (treatises, etc.)	**STEP 4 — Identifying the correct/expected output** In this phase, the specific situation is analyzed, taking into consideration the data that the AI, theoretically, should have had available (input it should have received). In analogy to the classical reconstruction of ideal medical conduct, through the analysis of scientific sources, such as guidelines (national and international), consensus documents (national and international), operating procedures (local, national and international), evidence-based publications (Cochrane reviews, *meta*-analyses, etc.) and other literature data, consisting of treatises and articles published in peer-reviewed journals (PubMed-Medline, Embase, Scopus, Ovid, ISI Web of Science, etc.), the correct/expected output is identified. These scientific sources of non-equivalent importance are to be classified according to the source hierarchy below: – Guidelines – Consensus documents – Operational procedures – Evidence-based publications – National literature (treaties, etc.). This approach allows the correct expected output to be determined.
STEP 5 — Reconstruction of the real medical conduct After examining the sources and the ideal medical conduct, as described above in STEP 4, the medico-legal expert must establish whether there are sufficient data to proceed to the reconstruction and ascertainment of the conduct of medical and healthcare personnel. If this is not possible (i.e. salient data missing, incomplete documentation, lack of physiopathological links of pathological features, etc.), further ascertainment of possible Medical Responsibility and/or Liability cases. A formal conclusive report to elucidate why the evaluating process has been interrupted should be compiled.	**STEP 5 — Verifying the possibility of reconstructing the AI's decision-making process** It is necessary to assess if the AI model in question is an "*explainable AI*" (XAI) model or a more complex "*black box*" model. Once the actual existence of an error has been confirmed and the possibility of explaining it has been verified, it will be possible to start the reactive phase of the process.
STEP 6 — Reconstruction and verification of real conduct of medical and healthcare personnel The first phase consists of applying the extrapolation method to data, which are significant and useful for reconstructing and ascertaining the conduct of medical and healthcare personnel. Evaluation of the correctness of the various diagnostic, prognostic and therapeutic phases is carried out by comparing ideal conduct, desumed from referenced scientific sources, such as guidelines, consensus documents, operational procedures and evidence-based publications. However, the absence of a guideline is not *prima facie* evidence of negligence.	**STEP 6 — Verification of received inputs** The purpose of this step is to assess whether the inputs received by the model, which led to the erroneous output in question, were correct and appropriate for the specific task required. In this case, the error could be human in nature and not dependent on the performance of the system itself. Therefore, if an error is identified at this level, one can proceed directly to step 8, for the classification of the error. If, on the other hand, the input is judged to be correct and adequate for the task the AI was intended to perform, the causes of the error will be sought through a re-evaluation of performance.
STEP 7 — Identification of error/non-observance The process of analysis and comparative evaluation between ideal conduct and true conduct leads to the identification of possible error and/or non-observance of required rules of conduct. Type of error: – Real error: this is a material error, of omission or commission, due to violation of a universal and/or epidemiological scientific law or of consolidated rules of experience and competence; – Pseudo-error (apparent error): this is only an apparent error due to a general absence of scientific knowledge on a specific issue at the time of the event or, alternatively, related to an unpredictable and inevitable event (i.e. force majeure); – Conscious error: aware of having not identified the true (aetiology of the) pathological state of the patient, the medical doctor applies diagnostic or therapeutic procedures with only an *ex adiuvantibus* aim causing damage to that patient; – Non-observance of required rules of professional medical conduct;	**STEP 7 — Reassessment of model performance and identification of the source of the error** This step represents an accurate evaluation of the machine and its performance in the field, outside the "protected" context of the initial development phases, answering the question *"how does the model perform today?"*. In this context, two fundamental questions are asked: – Are the parameters obtained from the re-evaluation in accordance with those obtained during the development phase? – Was there adherence to the overall scientific truth, both in the development and implementation phases? If substantial changes are found, such as a decrease in performance, it is plausible to identify a *structural or intrinsic error in the model.* If, at the time the error occurred, the model no longer conforms to scientific truth (e.g. the dataset used does not appear adequate or best implementation practices have not been adopted), this may indicate the presence of a *structural or intrinsic error*.

(*continued on next page*)

Table 1 (*continued*)

Original Guidelines	Guidelines applied to AI
– No-fault medical accident.	If both above conditions are met, and despite this validity an error has occurred, the latter may have occurred due to the exceptionality of the case, thus being *an exception*.
STEP 8 — Classification of error/non-observance If the comparative evaluation between ideal conduct as desumed from scientific sources and real conduct reveals EVIDENCE of error(s) or non-observance of required rules of conduct, qualification-correlation of such error/non-observance (single or multiple) must be carried out, according to the specific area of expertise, as regards patient's consent and diagnostic, prognostic or therapeutic phase.	**STEP 8 — Error Classification** Once the error has been identified through the previous steps, it is essential to proceed to its classification. It is necessary to determine which step in the validation process revealed a deficiency in relation to the criteria considered in the previous steps. We can distinguish: – low accuracy – underfitting or overfitting – robustness – low specificity/sensitivity ● Input error - Frontend operator-dependent - Backend operator-dependent ● Structural/intrinsic error • Exceptional case
STEP 9 — Error evaluation—ex-ante. Possible causes of justification This evaluation involves the reasons for identified and classified error and/or non-observance. In particular, the medico-legal expert must establish whether the reasons for any such error and/or non-observance are true or whether there is a cause for justification (justifiable error). This evaluation phase requires the medico-legal expert to enter a state of ex-ante evaluation/judgement. Ex-ante evaluation must consider all (and only) the diagnostic, prognostic and therapeutic hypotheses which could be formulated a priori with respect to knowledge of the true pathological state/condition, desumed ex-post from the data collected after the event in question, since only such an evaluation can reflect the aspects of evaluation and decision making existing in the space–time conditions in which the medical and healthcare personnel were working, and their conduct as examined in those conditions. The medico-legal expert must supply technical reasons for cases of justifiable error.	**STEP 9 — Ex-ante error assessment. Possible causes of justification** The aim of this assessment is to identify any elements that may represent mitigating factors in the error and to define the reasons that led to the wrong output. It is exquisitely the responsibility of the medico-legal expert, who will enlist the help of specialists. It is up to the medico-legal expert to determine if the reasons for the errors and/or non-compliance are indeed methodologically reprehensible or if there is a justifiable cause (justifiable error). This assessment phase requires an ex-ante assessment/judgement approach. This phase appears to be of obvious difficulty if it is not a human but an instrument/algorithm that commits the error.
STEP 10 — Causal value and causal link between error and event The causal value and the relationship of an actual causal link must be evaluated by means of a "criterion of scientific probability", such as universal law, statistical law or criterion of rational credibility. If this is not possible, due to the absence of "explanatory laws", evaluation must be interrupted.	**STEP 10 — Causal link between identified error and incorrect output** At this point, after having identified the erroneous output produced by the model (step 3) and having identified its possible cause – the possible erroneous inputs (step 6) or the structural/intrinsic problems of the model itself (step 7) – to fully understand the relationship between them, it is necessary to establish the existence of the causal link. Causal value and the existence of an effective causal link must be assessed through a "scientific probability criterion", based on scientific laws, statistical laws and rational credibility criteria; Through this rigorous evaluation, an attempt will be made to establish the existence of a strong causal link between the error found and structural or intrinsic problems in the model. This process contributes to a deeper understanding of the underlying causes of the error and the identification of possible areas for improvement of the model itself.
STEP 11 — Universal law, statistical law or criterion of rational credibility The causal value of error and the relationship of an actual causal link between error/non-observance and damage may be evaluated according to: (a) universal laws, by means of deduction; (b) statistical laws, by means of inference; or, in the absence of such laws, according to (c) a criterion of rational credibility, i.e. referring only to the average experience and expertise of the medical category or class in question.	**STEP 11 — Universal law, statistical law or rational credibility criterion** The causal value of the error and the relation of an effective causal link between error/incompetence and damage can be assessed according to: (a) universal laws, by deduction; (b) statistical laws, by inference; or, in the absence of such laws, according to (c) a rational credibility criterion, i.e. by referring only to the experience and average competence of the medical category or class in question. It seems clear, however, that in a context involving the use of AI systems, the analysis of such laws and criteria will require the assistance of new professional experts who will assist the medico-legal expert in assessing the individual case.
STEP 12 — Identification of the degree of probability of causal value and causal link A later check of the causal value and causal link between error and injury must be made by applying counterfactual reasoning and eventually additional criteria. The conclusion must be expressed in terms of near certainty, probability (when possible, estimating the percentage of probability) or exclusion of the causal value–causal link between error/non-observance and damage.	**STEP 12 — Identifying the degree of probability of causal value and causal link** A subsequent verification of the causal value and causal link between the error and the erroneous output must be carried out by applying counterfactual reasoning and any further criteria. This can lead to excluding the causal link or reinforcing it.
STEP 13 — Damage estimation At the end of medico-legal evaluation, whether within the juridical ambit or outside it, the medico-legal expert must quantify the temporary or permanent biological injury causally correlated with error/non-observance.	**STEP 13 — Damage estimation and model improvement (machine learning)** The ultimate goal of the whole process is the estimation of the damage and, from a risk management perspective, the improvement of the model itself (machine learning).

reconstruction of the sequence leading from the misclassification to the adverse outcome. Finally, damage quantification and recommendations for algorithm improvement (e.g., re-training on more representative datasets, adjustment of thresholds for high-risk profiles) would complete the 13-step process. This simplified case scenario exemplifies how the proposed methodology can structure medico-legal reasoning in AI-related malpractice cases.

Despite these strengths, several limitations must be acknowledged. The framework is conceptual and requires validation through empirical studies, including real-world cases of AI-related adverse events. Furthermore, while grounded in European forensic tradition [23], medico-legal reasoning varies across jurisdictions. Civil law systems may be more receptive to structured methodological guidelines, while common law systems may require integration with case precedent and adversarial testing [24]. The framework must therefore be adapted to local legal contexts to ensure international applicability.

From a medico-legal training and certification perspective, the spread of AI in healthcare calls for structured pathways specifically targeting experts involved in AI-related claims. Multidisciplinary programmes (e.g., postgraduate courses or university Masters) including

Table 2
Detailed information about the adaptation of the original European Council of Legal Medicine (ECLM) Guidelines to Artificial Intelligence in healthcare.

Step 1 **Comparative assessment of available AI data**	The first phase involves an evaluation of the data available to the AI. The so-called dataset. A machine learning dataset is a structured set of data used for training and evaluating machine learning algorithms. A good training dataset should be large, diversified and balanced, containing a wide range of examples that accurately represent all the different situations and variables that may arise during implementation. Particular attention should be paid to biases and potential distortion in the collected data. In general, this bias may be a direct consequence of a conscious or unconscious sampling error during data collection. It may be caused by technical-operational factors (e.g. instrumental limitations, such as the resolution, accuracy and dynamic range of the equipment, both medical and non-medical, used for data collection) or by socio-cultural factors (e.g. due to under-represented population segments in existing data collection campaigns or because of unconscious biases or prejudices of the operators responsible for data collection and processing). It should be noted that these drifts, already present upstream in the dataset prior to any AI development phase, ooze by statistical effect within the algorithms developed from these datasets, potentially turning into output errors. The construction of a good dataset implies an integrated and multidisciplinary approach that cannot disregard data analysis: The collection and use of data cannot disregard the applicable laws, both at international (European for European countries) and national level. Evidence of this is the fact that, in March 2023, due to ChatGPT's non-compliance with data processing consent rules, the Italian Privacy Guarantor opened an investigation against OpenAI, which led to the suspension of the service in Italy. This highlights how, even in the presence of a well-structured and technically sound model, its implementation in some territories is marred by non-compliance with local regulations. A domain expert is a person who possesses in-depth knowledge and understanding of the area in which the model is to be applied, but who has no specific expertise in AI design (13). In the medical field, this would be a specialist in the medical matter in question who also has documented experience in clinical studies and thus knowledge of the indispensable data to be collected for the construction of the dataset. In general, unlike many clinical studies conducted so far, the construction of a good dataset will require the heterogeneity of the data in terms of gender, geographic origin (also considering differences in diet and habits), age (unless referring to specific age groups), etc. At this stage a discussion with medico-legal experts would certainly be of fundamental help to avoid errors in the selection of data for the construction of the dataset. In this case, it is the technician (e.g. a bioengineer or an engineer with similar skills) who has specific technical knowledge of database systems and AI algorithms applied to them in the biomedical field. He is the one who builds the database. He does this on the one hand based on the specialist's expertise, and on the other hand considering the legal requirements, obviously adding his knowledge and skills on how the model itself works. In particular, it will be the task of the data collector, in synergy with the medical specialist (domain expert), to collect as many clinical features as possible (per patient, possibly with follow-up and in the various possible formats: images, videos, signals, tabular data from medical records, laboratory tests, omics data) known to have a potential correlation with the pathology or condition under investigation. Indeed, in an AI algorithm it is not possible to know in advance, before its design, which features will be most relevant in its design. This information will be highlighted during the development and testing of the model itself. It will therefore be crucial to have as large a data set as possible. Such data must not be characterized by biases of any kind but represent the conditions that exist in the real world in terms of probability distributions of continuous features, frequency of categorical features, frequency of events, variability of images and signals, reflected in the various possible physio-pathological conditions of the individual. Once the data has been collected, it will be the task of the *data collector* to homogenize it and transform it into a structured set that can be used by the model, but which can also represent a complete and easily accessible database for future checks. These transformations are part of the technical phase of data *pre-processing*, which includes, for example, the conversion of variables reported in different units of measurement into the same unit of measurement, the conversion of data reported in text format into a numerical format interpretable by the algorithm, and the eventual normalization of the data necessary for some specific AI algorithms, such as KNN (k-nearest neighbors algorithm). It is essential, for medico-legal purposes, that all transformations reported on the data are documented and archived, to be easily accessible during future audits. The *data collector* uses, both in the preparation of the data set as well as in the training (*training*) and evaluation (*testing*) phase, his own expertise on the technical functioning of the model based on the state of the art deduced from his own knowledge and experience as well as from *peer-reviewed* scientific publications relevant at the time of algorithm development. It must also be aware that choices made during data collection and preliminary processing (e.g. mathematical transformations of data, inclusion/exclusion of samples and/or variables) must be made in a rationally motivated manner and in accordance with industry *best practice* (see step 2). This is because the choices have a direct and not always easily predictable effect on the algorithm's decision-making mechanism (by influencing its training) and are, therefore, a potential channel for erroneous evaluations of results. Finally, it is important to remember that the choice of the specific clinical use case and the type of data collected will influence the final model architecture. This phase is to be considered preliminary to the development of the model. The model must therefore first comply with the above. It is therefore necessary for several professionals to be involved in the development to ensure compliance with the legal system of reference and the needs of the specialist who will use it. To date, in fact, little collaboration between back-end operators (developers) and front-end operators (end users) in the construction of models (12) has emerged. ■ *National and supranational laws* ■ *Domain Expert* ■ *Data Collector*
Step 2 **Validation of the dataset based on scientific evidence**	The second step involves validating the substance of the dataset based on so-called *scientific references*. These include guidelines, best practices and scientific reference literature. These are used to assess the scientific assumptions on which the model is based, such as the statistically significant size of the data set, minimization of bias (e.g. gender, ethnicity, age, anamnestic details). Verifying the statistically significant size of the data set is essential to ensure the representativeness of the data used to train the model. Bias minimization is equally crucial, as it helps to guarantee that the model does not develop unwanted bias or discrimination. Finally, performance evaluation, measured in terms of specificity and sensitivity, provides a clear overview of the effectiveness of the model. Also in this second phase, the comparison with the forensic scientist can be useful as an input to assess the consistency (*continued on next page*)

Table 2 (*continued*)

	and reliability of the model in specific contexts. A model that does not satisfy the basic criteria for its development, as indicated in step 1, or that does not follow the scientific evidence during the development process, as indicated in step 2, would present a significant defect from the outset. In such circumstances, the model should not reach the market, as it does not meet the essential requirements to ensure its reliability and safety.
Step 3 **Identification of incorrect output**	Step 3 considers the hypothesis of an AI result that turns out to be incorrect. The purpose of this phase is to precisely identify the erroneous output and clarify its nature (e.g. false positive, false negative, error in therapy, incorrect prognosis, error in follow-up, etc.). This step provides a precise understanding of the type (nature) of the error and the context in which it occurred.
Step 4 **Identifying the correct/expected output**	In this phase, the specific situation is analyzed by carefully examining the data the AI should ideally have possessed or the input data it should have received. This is a critical evaluation aimed at comparing the reality of the data actually available with the expectations of the information that the AI should have used to generate an accurate output. This process aims to identify any discrepancies between what the AI theoretically had available and the actual situation under consideration. In this way, by carefully analyzing the expected inputs and those actually received, it is possible to identify any shortcomings, errors or limitations in data collection, thus contributing to a better understanding of the underlying causes of any discrepancy between the desired output and that actually produced by the AI. This approach makes it possible to determine what the correct expected output would have been. This is an analogy to the classical reconstruction of ideal medical conduct from case information. This ideal conduct is reconstructed by analyzing scientific sources, such as guidelines (national and international), consensus documents (national and international), operating procedures (local, national and international), evidence-based publications (Cochrane reviews, *meta*-analyses, etc.) and other literature data, consisting of treatises and articles published in peer-reviewed journals (PubMed-Medline, Embase, Scopus, Ovid, ISI Web of Science, etc.), preferably with an impact factor.
Step 5 **Verifying the possibility of reconstructing the AI's decision-making process**	Before proceeding to reconstruct the chain of events that led to the error, it is crucial to establish if this is possible or not. It is therefore necessary to assess whether the AI model in question is an *explainable AI* model or a more complex model, referred to as a *"black box"*. This assessment, made in consultation with the professionals involved in algorithm development, depends largely on the technical and algorithmic characteristics of AI. In particular, some families of methods are inherently explainable, since they are the result of the combination of simple mathematical rules applied to the variables provided as input (e.g. "tree" algorithms). In other cases, e.g. *deep neural networks*, the interpretation and rationalization of the algorithm's decision-making process is less linear and must therefore be dealt with on a case-by-case basis. More specifically, in comparison with the clinical context, one possible method for defining the clinical question of interest and identifying the processes for which evidence needs to be collected and evaluated is the construction of models or causal pathways. A causal pathway is a diagram that illustrates the links between the intervention(s) of interest and the intermediate, surrogate and health outcomes that are thought to influence the intervention(s) (which we might liken to legally relevant concauses). In designing the pathway, guideline developers make explicit the premises on which their hypotheses of effectiveness are based and the outcomes (benefits and harms) they consider important. This identifies the specific questions that the evidence must answer to justify the efficacy conclusions and highlights gaps in the evidence, for which future research is needed. Turning to the context of AI, from a historical perspective, the first clinical decision support systems, developed since the 1950 s, were rule-based systems that attempted to mimic human experience in specific medical fields. However, such systems, which can be compared to guidelines, are static, rigid and do not analyze the patient's state with a personalized approach. Over the years, thanks also to the advancement of computing power and the availability of *big data*, together with the ever more pressing demands of precision medicine, new algorithms have been devised that are increasingly complex and inexplicable, yet increasingly high-performance (such as the aforementioned deep neural networks). We were faced with the well-known *trade-off* (crossroads) between accuracy and transparency in the technical field. Despite the first rule-based systems being fully explicable, tractable in one sense and the other (i.e. allowing the reason for a particular operational choice to be derived a posteriori), current scientific production supports the use of more complex and opaque systems, with the introduction of new simplification methods that maintain the same or a slightly lower level of accuracy, flexibility and customization. This paradigm shift was made possible by explainable AI, the subject that studies the degree of transparency of AI models and possible approaches to improve understanding of opaque and complex AI models. *Explainable AI*, also known as XAI (Explainable Artificial Intelligence), is characterized by the ability to understand and explain decisions made by its AI system clearly and understandable to humans. This enables developers, users and stakeholders to understand how AI models reach certain conclusions and, in particular, to understand the rational and mathematical link between input data and the final output. In the case of XAI, the decision-making process is – by definition – explainable and allows a precise and detailed reconstruction of the motivations and numerical rules that determined the output of the model. Consequently, in the case of XAI, the subsequent steps will certainly be simpler and based on the information provided by the algorithm itself. In the case, however, of a *black box* algorithmic system, we refer to a system or algorithm that produces results without providing a clear understanding of how they were obtained. In other words, the input and output of the system are known, but the internal process that connects them remains opaque and difficult to interpret. In the case of a black box, one can hypothesize the possibility of indirectly reconstructing these mechanisms by re-analyzing the system's performance, by specialized technical personnel, to obtain an evaluation of the metrics as close as possible to the time when the erroneous output was generated. Compared to explainable models, reconstruction in this situation will certainly be a more complex and costly process, requiring a global re-evaluation of the entire model (see below). Much of the explanatory process is done by following a backward approach, i.e. by repeatedly observing the outputs produced by the algorithm in response to appropriately calibrated inputs, including inputs like those that generated the error under analysis. This makes it possible to analyze the sensitivity and stability of the black box with respect to specific groups of variables and ranges of values, thus enabling an indirect explanation of the model's decision-making processes. Furthermore, the effect of errors (of measurement, transcription, etc.) of the input variables on the output of the algorithm, i.e. the propagation of the error through the different levels (*layer*) of the AI, up to its final decision, will be worthy of particular evaluation due to its relevance in the context of the explainability of the output. Once the existence of an error has been confirmed and the possibility of explaining it has been verified, it will be possible to start the actual reactive phase of the process that, starting from the adverse event, reconstructs backwards the sequence of events with the aim of identifying the factors that caused or contributed to the occurrence of the event itself. Through an accurate methodology, the erroneous output, its cause and the causal link between the initial conditions, the consequent reasoning and its effects will be assessed.

(*continued on next page*)

Table 2 (*continued*)

Step 6 **Verification of received inputs**	The objective of this phase is to verify the inputs supplied to the model that led to the production of the erroneous output in question. An attempt is made to determine if the inputs themselves were correct and adequate for the specific task. It is crucial to consider the possibility that the inputs responsible for the system's erroneous response are inherently inconsistent. In such a case, the error may be attributed to a human component and not necessarily to the performance of the system itself. Therefore, if an error is identified at this level, it is possible to proceed directly to step 8 – a step in which errors are classified – without having to initiate a complex and time-consuming process of re-evaluating the system, as described in the next step (step 7). In the case where the inputs are judged to be correct and suitable for the task assigned to the AI, the causes of the error are investigated through a re-evaluation of the performance of the model itself, to obtain a detailed overview of the system's performance also with the aim of identifying possible areas of improvement to optimize its accuracy and reliability.
Step 7 **Reassessment of model performance and identification of the source of the error**	The verification of the actual performance of the AI, in this specific case, cannot disregard the exact knowledge of the development point – independent of the human hand – at which, thanks to machine learning (ML), the model itself had arrived. The characteristics and performance of the machine may, by the nature of the algorithms themselves, vary from the initial design stages: so many changes, even if predictable, may accumulate over time and generate an unexpected divergence in the final performance of the AI/ML-based software. We are referring to the natural evolution of algorithms that, as they process more and more data, are enriched with information that leads to results that, in earlier times, would not have been evident. This step represents an accurate assessment of the machine and its performance in the field, i.e. outside the 'protected' context of the initial development phases, answering the question *"how does the model perform today?"*. This passage opens a parenthesis on the concept of "*life-cycle assessment*", a proposal currently much discussed in the literature on AI regulation. This concept provides that, since AI products by their nature change over time, re-evaluation procedures should be put in place even after they have been placed on the market. As far as the evaluation of model performance is concerned, there is currently no established consensus in the scientific literature on how it should be conducted; it is likely that in the future new scientific evidence will provide more specificity on this aspect and new emerging professionals will enable its implementation. However, it is possible to take a cue from a rather well-structured proposal, the CLAIM (Check List for Artificial Intelligence in Medical Imaging) guidelines, which, regarding the evaluation of a model, consider the following parameters – Performance metrics – e.g. measurement of prediction accuracy – Statistical significance – e.g. the definition of confidence intervals – Robustness of the model – e.g. measurement of sensitivity and specificity – New tests – repeating the model test on new datasets other than the training dataset. Currently, laboratories providing data from technical equipment, such as toxicology or genetics laboratories, are subject to periodic renewals of quality certifications. In analogy to this, one could imagine a periodic reassessment of the performance of the AI model. This would make it possible both to contain the risk of erroneous outputs and to reconstruct more accurately the characteristics of the model at the time when the error occurred. Once the actual performance of the machine, its characteristics and any criticalities have been defined, it is possible to compare them with those obtained in the development phase. In this context, two fundamental questions arise: – Are the parameters obtained from the re-evaluation in accordance with those obtained in the development phase? – Was there adherence to the overall scientific truth, both in the development and implementation phases? The first question to ask is whether the machine's current performance corresponds to the initial performance when the model was released on the market. If there are substantial variations, such as a decrease in performance, it is plausible to identify a structural or intrinsic error in the model. The second question focuses on conformity and adherence to the overall scientific truth. It investigates the same aspects as defined in step 2, i.e. whether the model conforms to guidelines, best practices and relevant scientific literature. Differently from the proactive phase, which only considers the development phase, in this case the model is considered at the time when the error occurred. If, at the time the error occurred, the model no longer conforms to scientific truth (e.g. the dataset used does not seem adequate or best implementation practices have not been adopted), this may indicate the presence of a structural or intrinsic problem. As mentioned earlier, this type of error could be mitigated by periodic re-evaluations of the model. If both above conditions are met, i.e. the actual performance is in line and there is adherence to the general scientific truth, the model can be considered valid in terms of performance and scientific correctness. However, if despite validity an error has occurred, this may be due to the exceptionality of the case. It will be carefully considered during the model implementation process, but the error cannot be attributed to a structural defect in the model, which, being an exception, could not have predicted it.
Step 8 **Error classification**	Once the error has been identified through the previous steps, it is essential to proceed to its classification. At this point, it is necessary to determine which step in the validation process revealed a deficiency in relation to the criteria considered in the previous steps. It is therefore possible to distinguish: – Input error – Structural/intrinsic error – Exceptionality of the case Each of these categories can in turn be further subdivided to obtain as specific a definition as possible of the type of error: – low accuracy – underfitting or overfitting – robustness – low specificity/sensitivity Input errors can be divided into several categories. The grossest error is the material errors in data input by frontend operator. More complex is the input error by backend operator. This error depends on an incorrect or deficient design of the platform. It could be, for instance, an ineffective antivirus with data theft or misleading automatic data entry. The platform might have been designed with too high degree of freedom for user input, or it might contain data handling errors that affect the functioning of the underlying artificial intelligence model. Structural errors of the model must be analyzed and framed from a technical point of view and classified according to their specific categories. Examples include: The accurate identification of such technical errors is crucial to identify the specific areas where the model has shown shortcomings. This is important not only as a basis for defining responsibilities, but above all to enable improvement

(*continued on next page*)

Table 2 (*continued*)

	measures to be implemented. Through this optimization process, which is the last stage of the proposed model, machine learning algorithms can learn and improve their performance to avoid repeating the same errors in the future. ● Input error - Frontend operator-dependent - Backend operator-dependent ● Structural/intrinsic error • For performance metrics, low accuracy might arise; • For statistical aspects, issues related to inference, such as underfitting or overfitting, might arise; • Regarding robustness, it might become apparent that the dataset used is too limited; • The model could produce a high number of false negatives, highlighting a sensitivity problem.
Step 9 **Ex-ante error assessment. Possible causes of justification**	The aim of this assessment is to identify any elements that may represent mitigating factors of the error and to define the reasons that led to an erroneous result. This is exquisitely the responsibility of the medico-legal expert, who will be assisted by specialists. In particular, it is up to the medico-legal expert to establish if the reasons of the errors and/or non-conformities are reprehensible from a methodological point of view or if there is a justifiable cause (justifiable error). This assessment step requires the medico-legal expert to adopt an ex-ante evaluation/judgement approach.
Step 10 **Causal link between identified error and incorrect output**	At this point, after having identified the erroneous output produced by the model (step 3) and having identified its possible cause – the possible erroneous inputs (step 6) or the structural/intrinsic problems of the model itself (step 7) – to fully understand the relationship between them, it is necessary to establish the existence of the causal link. Causal value and the existence of an effective causal link must be assessed through a "scientific probability criterion", based on scientific laws, statistical laws and rational credibility criteria; these are the three fundamental points on which medico-legal methodology is based. Considering what has emerged in the previous steps, this evaluation implies an in-depth analysis that goes beyond the simple observation of the wrong output but aims to causally link it with the error identified in the specific case. It will be crucial to consider how scientific laws and statistical knowledge apply to the specific context and to assess the validity of the assumptions and procedures adopted in the model. Furthermore, it will be necessary to check if the model has followed rational credibility criteria, such as the use of accepted and field-validated methods and approaches. Through this rigorous evaluation, an attempt will be made to establish the existence of a strong causal link between the error found and structural or intrinsic problems in the model. This process contributes to a deeper understanding of the underlying causes of the error and the identification of possible areas for improvement of the model itself.
Step 11 **Universal law, statistical law or rational credibility criterion**	The causal value of the error and the relation of an effective causal link between error/incompetence and damage can be assessed according to: a) universal laws, by deduction; b) statistical laws, by inference; or, in the absence of such laws, according to c) a rational credibility criterion, i.e. by referring only to the experience and average competence of the medical category or class in question. It seems clear, however, that in a context involving the use of AI systems, the analysis of such laws and criteria will require the assistance of new professional experts who will assist the medico-legal expert in assessing the individual case.
Step 12 **Identifying the degree of probability of causal value and causal link**	A subsequent verification of the causal value and causal link between the error and the incorrect output must be carried out by applying counterfactual reasoning. This is a mental process that consists of imagining and evaluating alternative situations to the one that happened. In other words, it involves considering what the outcome or event would have been like if the output had been correct. This can lead to excluding the causal link or reinforcing it.
Step 13 **Damage estimation and model improvement (machine learning)**	The ultimate goal of the entire process presented so far is damage estimation and, from a risk management perspective, the improvement of the model itself. For obvious reasons, we will not focus on damage estimation. About the possible improvement of the model, what has been reported so far shows how the errors identified play a fundamental role in allowing the algorithm to evolve and provide increasingly accurate output; based on the classification of errors carried out previously, precise optimization measures can be adopted. Optimization of machine learning is the process of continuously improving the accuracy of a model, aimed at increasingly minimizing the degree of error. With regard to resolving the structural errors defined above, specific interventions or even modification of the entire algorithm may be required. For example, if a low-precision problem is detected, optimization of the model would involve adjusting the weights to minimize the loss function, which is used to assess the accuracy of the predictions and make the model's predictions as close to correct as possible. This adjustment, known as "*fine-tuning*", requires high precision, as even small changes can have significant effects on the entire system. Exceptions, on the other hand, are theoretically less complex to correct. By providing the correct output in relation to an exceptional situation, the model will take this information into account in future predictions, providing an additional tool for resolving similar cases.

legal medicine, bioethics, AI regulation, and basic technical concepts could provide a common knowledge base. In parallel, continuing professional development modules and, where feasible, formal accreditation schemes for medico-legal experts dealing with AI-assisted care would help to standardise competencies and support courts and regulators in identifying suitably qualified experts.

Looking forward, the proposed methodology could serve as a building block for global standards in AI-related medical liability. By bridging medico-legal expertise with technical evaluation, it offers a practical tool for courts, regulators, and healthcare providers [25]. Future research should include pilot implementations of the 13-step framework within European healthcare systems and medico-legal services, for example as structured checklists or digital decision-support tools. Comparative studies across different European operational contexts—such as emergency care, radiology, pathology, and telemedicine—could assess feasibility, reproducibility, and impact on both error prevention and liability assessment. In addition, the progressive translation of the prompt-based structure into dedicated software platforms, interoperable with electronic health records and compliant with European data protection standards, represents a key avenue for transforming the conceptual framework into a practical, context-sensitive instrument for daily use. Such steps would contribute not only to fairer liability assessments but also to safer and more transparent integration of AI into healthcare systems worldwide.

5. Conclusions

AI is rapidly transforming healthcare, but its adoption raises unprecedented medico-legal questions. Existing liability frameworks, designed for human decision-making, often prove insufficient when applied to algorithmic errors and opaque models. This paper proposes an adaptation of a well-established forensic methodology that integrates proactive and reactive approaches to risk analysis, thereby offering a structured pathway for evaluating liability in AI-driven healthcare.

By aligning medico-legal reasoning with the specific characteristics of AI systems, the proposed framework supports clinicians, developers, and policymakers in preventing errors, reconstructing adverse events, and improving model safety. While further validation in real-world contexts is required, this methodology represents a step toward internationally applicable standards that can strengthen accountability and foster patient trust in the era of AI-assisted medicine.

6. Source of founding

None.

7. Disclaimers

The views expressed in the submitted article are our own and not an official position of the institution.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] O. Dedehayir, M. Steinert, The hype cycle model: a review and future directions, Technol. Forecast. Soc. Change 108 (2016) 28–41, https://doi.org/10.1016/j.techfore.2016.04.005.
[2] R. Worthington, The social control of technology. By David Collingridge, Am. Polit. Sci. Rev. 76 (1) (1982) 134–135. https://doi.org/10.2307/1960740.
[3] O. Kudina, P.P. Verbeek, Ethics from within: Google Glass, the Collingridge dilemma, and the mediated value of privacy, Sci. Technol. Hum. Values 44 (2) (2019) 291–314, https://doi.org/10.1177/0162243918793711.
[4] European Commission, Artificial Intelligence for Europe—Communication, Brussels, 25.4.2018 COM, 237 final, 2018. https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe.
[5] European Parliament and Council of the European Union, Regulation (EU) 2017/745 on medical devices, Off. J. Eur. Union L 117 (2017) 1–175. https://eur-lex.europa.eu/eli/reg/2017/745/oj.
[6] R. Cecchi, T.M. Haja, F. Calabrò, I. Fasterholdt, B.S.B. Rasmussen, Artificial intelligence in healthcare: why not apply the medico-legal method starting with the Collingridge dilemma? Int. J. Leg. Med. 138 (3) (2024) 1173–1178, https://doi.org/10.1007/s00414-023-03047-2.
[7] I.G. Cohen, T. Evgeniou, S. Gerke, T. Minssen, The European artificial intelligence strategy: implications and challenges for digital health, Lancet Digit. Health 2 (7) (2020) 376–379, https://doi.org/10.1016/S2589-7500(20)30105-X.
[8] European Commission, High-Level Expert Group on Artificial Intelligence, Ethics guidelines for trustworthy AI, 2019. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.
[9] European Commission, Proposal for a regulation on a European approach for artificial intelligence, COM(2021) 206 final, Brussels, 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206.
[10] S.D. Ferrara, E. Baccino, T. Bajanowski, R. Boscolo-Berto, M. Castellano, R. De Angel, A. Pauliukevičius, P. Ricci, P. Vanezis, D.N. Vieira, G. Viel, E. Villanueva, Malpractice and medical liability. European guidelines on methods of ascertainment and criteria of evaluation, Int. J. Legal Med. 127 (3) (2013) 545–557. https://doi.org/10.1007/s00414-013-0835-z.
[11] R. Cecchi, F. Calabrò, T.M. Haja, M. Sperti, E.A. Zizzi, M.A. Deriu, Intelligenza artificiale in sanità: proposta di una nuova metodologia medico-legale in materia di responsabilità medica, Riv. Ital. Med. Leg. Diritto Sanit. 3–4 (2024) 441–462. https://doi.org/10.36149/0390-1913-290.
[12] Brussels (2021). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206.
[13] U.s., Food and Drug Administration, Artificial intelligence and machine learning (AI/ML)-enabled medical devices: guidance for industry and Food and Drug Administration staff, FDA, Silver Spring, MD, 2022.
[14] World Health Organization, Ethics and governance of artificial intelligence for health: WHO guidance, WHO, Geneva, 2021 https://www.who.int/publications/i/item/9789240029200.
[15] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, A. Galstyan, A survey on bias and fairness in machine learning, ACM Comput. Surv. 54 (6) (2021) 1–35, https://doi.org/10.1145/3457607.
[16] Z. Obermeyer, B. Powers, C. Vogeli, S. Mullainathan, Dissecting racial bias in an algorithm used to manage the health of populations, Science 366 (6464) (2019) 447–453, https://doi.org/10.1126/science.aax2342.
[17] W.N. Price, I.G. Cohen, Privacy in the age of medical big data, Nat. Med. 25 (2019) 37–43, https://doi.org/10.1038/s41591-018-0272-7.
[18] F. Doshi-Velez, B. Kim, Towards a rigorous science of interpretable machine learning, arXiv preprint arXiv:1702.08608 (2017). https://doi.org/10.48550/arXiv.1702.08608.
[19] C. Rudin, Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, Nat. Mach. Intell. 1 (5) (2019) 206–215, https://doi.org/10.1038/s42256-019-0048-x.
[20] R. Cecchi, D. Cusack, B. Ludes, B. Madea, D.N. Vieira, E. Keller, J. Payne-James, A. Sajantila, M. Vali, R. Zoia, N. Cucurachi, M.L. Schirripa, F. Marezza, L. Anzillotti, L. Donato, C. Cattaneo, D. Favretto, S. Pelotti, V. Pinchi, S. Vanin, M. Gherardi, European Council of Legal Medicine (ECLM) on-site inspection forms for forensic pathology, anthropology, odontology, genetics, entomology and toxicology for forensic and medico-legal scene and corpse investigation: the Parma form, Int. J. Leg. Med. 136 (4) (2022) 1037–1049, https://doi.org/10.1007/s00414-021-02734-5.
[21] D. Cusack, S.D. Ferrara, E. Keller, B. Ludes, P. Mangin, M. Väli, N. Vieira, European Council of Legal Medicine (ECLM) principles for on-site forensic and medico-legal scene and corpse investigation, Int. J. Leg. Med. 131 (4) (2017) 1119–1122, https://doi.org/10.1007/s00414-016-1479-0.
[22] J. Camatti, A.L. Santunione, M. Bolognini, D. Cusack, S. Zerbo, A. Argo, M. Puntarello, G. Scalzo, P. Fattorini, T. Bruscagin, L. Desinan, L. Battistig, F. Tosku, G. Viel, G. Franchetti, B. Solarino, L. Ambrosi, G. Cecchetto, S.D. Visonà, E. Sala, A. Oliva, G. Mercuri, E.S. Oliveri, P. Fais, S. Lippi, S. Bianchini, U. Genovese, L. Franceschetti, A. Asmundo, E. Ventura Spagnolo, G. Burrascano, A. Verzeletti, F. Attico, V. Sarallo, I. Aquila, M.A. Sacco, S. Gualtieri, G. Mammola, A. Attanasi, D.M. Pingaro, V. Pinchi, M. Focardi, F. Ventura, I. Caristo, N. Vernazza, C. Casella, G. Di Donna, M. Marisei, C.P. Campobasso, A. Feola, P. Palermo, V. Bugelli, K. Doci, M. Picozzi, M. Burlando, F. Vecchio, M. Gabbrielli, F. Baldari, R. Cecchi, Towards a standard of scientific evidence in on-site inspection: compilation of the ECLM on-site inspection form in a broad case history, Leg. Med. (Tokyo) 78 (2025) 102717, https://doi.org/10.1016/j.legalmed.2025.102717.
[23] European Council of Legal Medicine, Guidelines for the assessment of medical liability, ECLM, Strasbourg, 2014.
[24] J. Henderson, R. Wallace, S. Wolf, Artificial intelligence in health care: balancing innovation and regulation, J. Law Med. Ethics 50 (1) (2022) 44–52, https://doi.org/10.1017/jme.2022.8.
[25] L. Floridi, M. Holweg, M. Taddeo, AI for social good: unlocking the opportunity, Sci. Public Policy 49 (1) (2022) 1–11. https://doi.org/10.1093/scipol/scab066. Sci Public Policy 49(1):1–11. https://doi.org/10.1093/scipol/scab066.